



Encouraging a data citation culture

Daniel Noesgaard, GBIF Secretariat

17 OCT 2018, GB25, KILKENNY, IRELAND

WHAT ARE DATA CITATIONS?



A WEEKLY ILLUSTRATED JOURNAL OF SCIENCE

"To the talks ground Of Nature trusts the mind which builds for aye."-WOEDSWORTH

THURSDAY, NOVEMBER 4, 1869

NATURE: APHORISMS BY GOETHE NATURE! We are surrounded and embraced by her : powerless to separate ourselves from her, and powerless to penetrate beyond her. Without asking, or warning, she snatches us up into

her circling dance, and whirls us on until we are tired, and drop from her arms. She is ever shaping new forms: what is

yet been ; what has been, comes not an, Everything is new, and yet nought but he old. We live in her midst a brow her not. She is

We constantly act upon her, and yet have no power over her.

The one thing she seems to aim at is Individuality; yet she cares nothing for individuals. She is always building up and destroying; but her workshop is inaccessible.

Her life is in her children; but where is the mother? She is the only artist ; working-up the most uniform material into utter opposites ; arriving, without a trace of effort, at perfection, at the most exact precision, though always veiled under a certain softness.

Each of her works has an essence of its own; each of her phenomena a special characterisation : tells them not whence they came, nor whither they and yet their diversity is in unity.

it herself, and yet she acts for us, the lookers-on. wear out, are always active and manifold.

Incessant life, development, and movement are in her, but she advances not. She changes for ever always renewing the spectators. Life is her most and ever, and rests not a moment. Quietude is exquisite invention ; and death is her expert coninconceivable to her, and she has laid her curse trivance to get plenty of life. upon rest. She is firm. Her steps are measured, her exceptions rare, her laws unchangeable.

She has always thought and always thinks ; though | earth, dull and heavy ; and yet is always shaking him not as a man, but as Nature. She broods over an until he attempts to soar above it.

all-comprehending idea, which to searching can find out Mankind dwell in h and she in them. With all men she plays a concord for love, and rejoices the more they win. We many, her moves are so hidden, that the game over before they know it.

which is most unnatural is still Nature ; the apidest philistinism has a touch of her genius. Whoso cannot see her everywhere, sees her nonever where rightly.

She loves herself, and her innumerable eyes and affections are fixed upon herself. She has divided herself that she may be her own delight. She incessantly speaking to perform the secret. causes an endless succession of new capacities for enjoyment to spring up, that her insatiable sympathy may be assuaged.

She rejoices in illusion. Whose destroys it in himself and others, him she punishes with the sternest tyranny. Whoso follows her in faith, him she takes as a child to her bosom.

Her children are numberless. To none is she altogether miserly; but she has her favourites, on whom she squanders much, and for whom she makes great sacrifices. Over greatness she spreads ber shield.

She tosses her creatures out of nothingness, and go. It is their business to run, she knows the road. She performs a play ; we know not whether she sees [] Her mechanism has few springs-but they never

The spectacle of Nature is always new, for she is

She wraps man in darkness, and makes him for ever long for light. She creates him dependent upon the 8.789.384 GEOREFERENCED RECORDS



"...data must be accorded due importance in the practice of scholarship and in the enduring scholarly record. In other words, data should be considered legitimate, citable products of research. Data citation, like the citation of other evidence and sources, is good research practice and is part of the scholarly ecosystem supporting data reuse."



WHY ARE DATA CITATIONS IMPORTANT?

Attribution

Rare marine fishes taken in Irish waters from 1786 to 2008	849
Seasearch Marine Surveys in Ireland	4k
Verified Marine records from Indicia-based surveys	4
Report on the Danish Oceanographical Expeditions 1908-1910 to the Mediter	27
Continuous Plankton Recorder Dataset (SAHFOS)	4k

Reproducibility

DOWNLOAD 2 OCTOBER 2018	
18,468 occurrences dow	nloaded
[D0] 10.15468/dl.gj74cn	
DOWNLOAD	
ILTER APPLIED 2 OCTOBER 2018	RERUN QUERY
Citation: GBIF.org (02 October 2018) GBIF Occurrence Download https://doi	.org/10.15468/dl.gj74cn
License: CC BY 4.0	
File: 3 MB Darwin Core Archive	
Make sure to read the data user agreement and citation guidelines.	
And	. A.P.
Country or area Ireland	
Gentific name Actinopterygil	



WHY ARE DATA CITATIONS IMPORTANT?

re	Tools	Inside (GBIF	-1-	Â×	٩	-	dnoesgaard
		0	DCCURRENCE DATASET	ı				
			Ireland's BioBlitz					
		F	Published by National Biodiversity Data Centre					
			Liam Lysaght ● ⊠ Barry O'Neill					
/ITY	≛ DOV	VNLOAD	⇔ HOME PAGE	1	43,163	2 0001	IRRENCE	30 CITATIONS
ed du	ring Irelar	nd's BioBl	itz N	letadata las	t mod	ified:	16 Janu	uary 2017
			D	ata last cha	naed:	15. Ja	nuary 2	017
				ata laot olla	ingea.	10.00	nuary 2	017
			н	losted by: N	ationa	l Biod	iversity	Data Centre
			H	losted by: N icense: CC I	ationa 3Y 4.0	l Biod	iversity	Data Centre



DATA CITATIONS IN A PERFECT WORLD

- Researcher downloads and uses data from GBIF.org
- Researcher publishes study and cites use of data by **DOI**, e.g.
- In methods: "We downloaded all records of ray-finned fishes in Ireland from GBIF¹⁴..."
- In references: 14. GBIF.org (02 October 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.gj74cn
- DOI reference is included in article metadata and deposited with **CrossRef**
- The link between paper and data is discoverable through the *Event* service
- We are notified of the event, i.e. the ink between a new publication and GBIFmediated data
- We (can) reciprocate the linking by updating de as it metadata deposited with DataCite
- The download now links to paper which links to the cownload...





DATA CITATIONS IN THE REAL WORLD





DOI CITATIONS IN 2018–THE GOOD NEWS





ENGAGING WITH AUTHORS





ENGAGING WITH JOURNAL PUBLISHERS





ENGAGING WITH JOURNAL PUBLISHERS...



Fuller, P. L., & Whelan, G. E. (2018). The flathead catfish invasion of the Great Lakes. Journal of Great Lakes Research. http://doi.org/10.1016/j.jglr.2018.07.001

Location data of flathead catfish occurrences were obtained from the Nonindigenous Aquatic Species (NAS) Database (Fuller and Nielson, 2015; U.S. Geological Survey (USGS) 2017a), the National Fish Habitat Partnership (2016), FishNet (http://fishnet2.net), VertNet (http:// vernet.org), and the Global Biodiversity Information Facility (GBIF) (http://gbif.org).



SCHOLIX

Author correspondence

Use of GBIF-mediated data in paper (10.

dnoesgaard Science Communications Coordinator

Sear Pam,

Thank you for using GBIF-mediated data in your recent publication: The flathead catfish invasion of the Great Lakes

I'm contacting you regarding the citation provided in this paper:

Global Biodiversity Information Facility (GBIF) (http://gbif.org)

However, our recommended standard citation for downloaded data is

· GBIF.org (date) GBIF Occurrence Download (DOI URL)

(note, the inclusion of the specific DOI allows for proper attribution of data reproducibility.)

Can you comment on why the recommended standard citation was not us data used in the paper, I would like to have it for our records.

DOIs in metadata



relation-type-id": "IsCitedBy", "related-identifier": "https://doi.org/10.1016/j.jglr.2

Links in Scholix

1 2 3 4 5

publication

The flathead catfish invasion of the Great Lakes Whelan Gary E., Fuller Pam L. Sources: Crossref Datasets in Datacite Publisher: Elsevier BV

Link to Datasets: 1



DATA USE HIGHLIGHTED

The flathead catfish invasion of the Great Lakes Pam L. Fuller * A B. Gary E. Whelan * B Show more

https://doi.org/10.1016/) /glr.201/4.07.501

Get rights and content

Abstract

A detailed review of historical literature and museum data revealed that flathead catfish were not historically native in the Great Lakes Basin, with the possible exception of a relict population in Lake Ene. The species has invaded Lake Ene, Lake St. Clair, Lake Huron, nearly all drainages in Michigan, and the Fox/Wolf and Milwaukee drainages in Wisconsin. They have not been collected from Lake Superior yet, and the temperature suitability of that lake is questionable. Flathead catfish have been stocked sparingly in the Great Lakes and is not the mechanism responsible for their spread. A stocking in 1968 in Ohio may be one exception to this, Dispersal resulted from both natural range expansions and unauthorized introductions. The invasion is ongoing, with the species invading both from the east and the west to meet in northern Lake Michigan. Much of this invasion has likely taken place since the 1990s. This species has been documented to have significant impacts on native fishes in other areas where it has been introduced; therefore, educating the public not to release them into new waters is important. Frequent monitoring of rivers and lakes for the presence of this species would detect new populations early so that management actions could be utilized on new populations if desired.

Keywords

Flathead catfish; Pylodictis olivaris; Great Lakes; Invasive





DATA USE HIGHLIGHTED





LINKS BETWEEN PAPERS AND DATA





• Enabling FAIR Data

- Coalition for Publishing Data in the Earth and Space Sciences (COPDESS)
- Convened by by American Geophysical Union (AGU)
- Funded by the Laura and John Arnold Foundation

• Align publishers and repositories in following best practices to enable **FAIR** and open data



FAIR is...

- Findable
- Accessible
- Interoperable
- Reusable





Statement of Commitment

- "...data must have unique and persistent identifiers and metadata appropriate to assist discovery and be cited in a form equivalent to other scholarly outputs..."
- "supplements will no longer be used as the primary archive for data"
- "using persistent identifiers in well-formed references to data"
- "Include a Data Availability Statement describing how the data underlying the findings of their paper can be accessed and reused"

Specifically, publishers are required to

- Review their journals' current guidelines to authors and editors
- Make sure that authors adhere to the FORCE11 Joint Declaration of Data Citation Principles



Signatories

- American Geophysical Union
- Copernicus Publications
- Ubiquity Press
- California Digital Library CDL

•<u>1</u> Wiley

- PANGAEA, Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI), Center for Marine Environmental Sciences, University of Bremen (MARUM)
- WDC Climate, Deutsches Klimarechenzentrum (DKRZ)
- Science and Science Advances
- •4 PLOS
- 2 Elsevier
- (Taylor & Francis)
- 3 (Springer Nature)

>65 per cent of all publications using GBIFmediated data



ENGAGING WITH JOURNAL PUBLISHERS

- FORCE2018 conference last week
 - The future of research communication and scholarship
 - Event data (Crossref <-> Datacite)
 - Standardized metrics (incl. DOI resolution)
- 1-on-1 with publishers
 - Elsevier
 - Wiley
 - PLOS
 - Taylor & Francis
 - SpringerNature
 - Sage
 - F1000
- Following up with citation statistics across journals



THANK YOU!

